

COMPITO DI APPLICAZIONI DI INTELLIGENZA ARTIFICIALE

15 Giugno 2010 (Punteggio su 30/30; Tempo 2h)

Esercizio 1 (punti 8)

Dato il seguente training set S:

Stazioni ferroviarie	Aeroporti	Raggiungibilità
2	1	No
1	2	Si
?	3	No
1	1	No
2	2	Si
1	3	No
2	1	Si
1	2	No
2	1	Si
1	3	No
2	1	Si
1	3	No
2	2	Si
1	2	No
?	3	Si

a) Si calcoli l'entropia del training set rispetto all'attributo **Raggiungibilità**

b) Si calcoli il guadagno dei due attributi rispetto a questi esempi di training

c) si costruisca un albero decisionale ad un solo livello per il training set dato, indicando le etichette delle foglie (numero di esempi finiti nella foglia/numero di esempi finiti nella foglia non appartenenti alla classe della foglia).

d) si classifichi l'istanza:

?	3
---	---

Esercizio 2 (punti 8)

L'esercizio su CLP si svolge alle ore 14 del 22 Giugno in Lab Info Grande.

Esercizio 3 (punti 8)

Nello stato iniziale descritto dalle seguenti formule atomiche:

**[at(paul,home), at(mark,home), at(shelf,home), at(table,home), at(fridge,home),
in(coke,fridge), in(phone, shelf), have_in_hand2(mark, book), have_in_hand1(paul, bottle),
hand_2_empty(paul), hand_1_empty(mark)]**

si vuole raggiungere il goal:

have_in_hand1(paul, phone), have_in_hand2(paul, book) have_in_hand1(mark, coke)

Le azioni sono modellate opportunamente come segue:

grab_with_hand1(Person,Item)

PRECOND: hand_1_empty(Person), at(Person,Location), in(Item,Box), at(Box,Location)

DELETE: hand_1_empty(Person), in(Item,Box)

ADD: have_in_hand1(Person,Item)

grab_with_hand2(Person,Item)

PRECOND: hand_2_empty(Person), at(Person,Location), in(Item,Box), at(Box,Location)

DELETE: hand_2_empty, in(Item,Box)

ADD: have_in_hand2(Person,Item)

release_from_hand1(Person,Item)

PRECOND: have_in_hand1(Person,Item)

DELETE: have_in_hand1(Person,Item)

ADD: hand_1_empty(Person), in(table,Item)

release_from_hand2(Person,Item)

PRECOND: have_in_hand2(Person,Item)

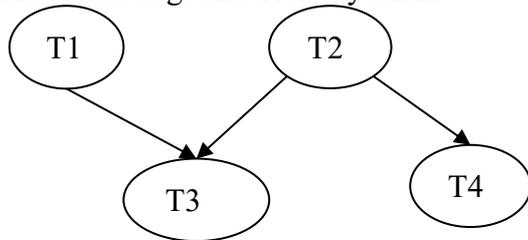
DELETE: have_in_hand2(Person,Item)

ADD: hand_2_empty(Person), in(table,Item)

Si risolva il problema utilizzando l'algoritmo POP. Si evidenzino i causal link e le minacce incontrate.

Esercizio 4 (punti 6) –per chi ha seguito nell'A.A. 2006/07 o successivi

Sia data la seguente rete bayesiana



Dove tutte le variabili assumono i valori vero e falso.

Le tabelle di probabilità condizionata sono

per T1:

	T1=Falso	T1=Vero
	0.4	0.6

per T2:

	T2=Falso	T2=Vero
	0.5	0.5

per T3:

T1	T2	T3=falso	T3=vero
Falso	Falso	0.2	0.8
Falso	Vero	0.3	0.7
Vero	Falso	0.7	0.3
Vero	Vero	0.7	0.3

per T4:

T2	T4=falso	T4=vero
Falso	0.2	0.8
Vero	0.3	0.7

Si calcoli la probabilità $P(T4|T1, \sim T3)$.

Esercizio 4a (punti 4) –per chi ha seguito nell'A.A. 2004/05 o precedenti

Spiegare brevemente le principali differenze tra forward e backward chaining e quale delle due è usualmente adottata nei sistemi di produzione.

SOLUZIONE

Esercizio 1

a) $\text{info}(S) = -7/15 * \log_2 7/15 - 8/15 * \log_2 8/15 = 0.997$

b)

Per calcolare il guadagno dell'attributo Stazioni non si usa l'entropia calcolata su tutto il training set ma solo sugli esempi che hanno Stazioni noto (insieme F):

$$\text{info}(F) = -6/13 * \log_2 6/13 - 7/13 * \log_2 7/13 = 0.996$$

$$\text{info}_{\text{Stazioni}}(F) = 7/13 * (-1/7 * \log_2 1/7 - 6/7 * \log_2 6/7) + 6/13 * (-5/6 * \log_2 5/6 - 1/6 * \log_2 1/6) = 0.538 * 0.591 + 0.462 * 0.650 = 0.618$$

$$\text{gain}(\text{Stazioni}) = 13/15 * (0.997 - 0.618) = 0.328$$

$$\text{splitinfo}(\text{Stazioni}) = -7/15 * \log_2(7/15) - 6/15 * \log_2(6/15) - 2/15 * \log_2(2/15) = 1.506$$

$$\text{gainratio}(\text{Stazioni}) = 0.328 / 1.506 = 0.218$$

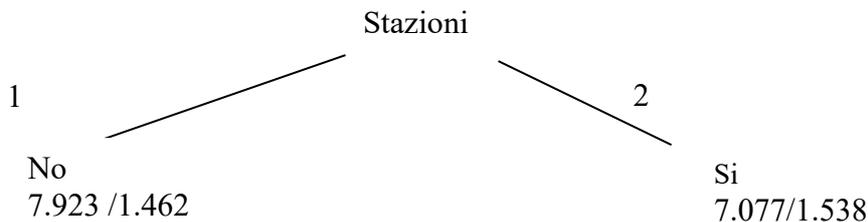
$$\text{info}_{\text{Aeroporti}}(S) = 5/15 * (-3/5 * \log_2 3/5 - 2/5 * \log_2 2/5) + 5/15 * (-3/5 * \log_2 3/5 - 2/5 * \log_2 2/5) + 5/15 * (-1/5 * \log_2 1/5 - 4/5 * \log_2 4/5) = 0.333 * 0.971 + 0.333 * 0.971 + 0.333 * 0.722 = 0.887$$

$$\text{gain}(\text{Aeroporti}) = 0.997 - 0.887 = 0.11$$

$$\text{splitinfo}(\text{Aeroporti}) = -5/15 * \log_2(5/15) - 5/15 * \log_2(5/15) - 5/15 * \log_2(5/15) = 1.585$$

$$\text{gainratio}(\text{Aeroporti}) = 0.11 / 1.585 = 0.067$$

c) L'attributo scelto per la radice dell'albero è Stazioni



d) l'istanza viene divisa in due parti, di peso rispettivamente 0.462 e 0.538. La prima parte viene mandata lungo il ramo 1 e classificata come No con probabilità $= 6.462 / 7.923 = 81.6\%$ e come Si con probabilità $= 1 - 0.816 = 18.4\%$. La seconda parte viene mandata lungo il ramo 2 e classificata come Si con probabilità $= 5.538 / 7.077 = 78.3\%$ e come No con probabilità $= 1 - 0.783 = 21.7\%$. Quindi in totale la classificazione dell'istanza è

$$\text{Si} = 0.462 * 18.4\% + 0.538 * 78.3\% = 50.6$$

$$\text{No} = 0.462 * 81.6\% + 0.538 * 21.7\% = 49.4$$

Esercizio 4 A.A. 2006/07 e successivi

$$P(T4|T1, \sim T3) = P(T4, T1, \sim T3) / P(T1, \sim T3)$$

$$P(T1, \sim T3, T4) = P(T1, T2, \sim T3, T4) + P(T1, \sim T2, \sim T3, T4)$$

$$P(T1, \sim T3) = P(T1, \sim T3, T4) + P(\sim T1, T2, \sim T3, T4) + P(\sim T1, \sim T2, \sim T3, T4)$$

$$P(T1, T2, \sim T3, T4) = P(T1)P(T2)P(\sim T3|T1, T2)P(T4|T2) = 0.6 * 0.5 * 0.7 * 0.7 = 0.147$$

$$P(T1, \sim T2, \sim T3, T4) = P(T1)P(\sim T2)P(\sim T3|T1, T2)P(T4|\sim T2) = 0.6 * 0.5 * 0.7 * 0.8 = 0.168$$

$$P(\sim T1, T2, \sim T3, T4) = P(\sim T1)P(T2)P(\sim T3|\sim T1, T2)P(T4|T2) = 0.4 * 0.5 * 0.3 * 0.7 = 0.042$$

$$P(\sim T_1, \sim T_2, \sim T_3, T_4) = P(\sim T_1)P(\sim T_2)P(\sim T_3 | \sim T_1, \sim T_2)P(T_4 | \sim T_2) = 0.4 * 0.5 * 0.2 * 0.8 = 0.032$$

$$P(T_1, \sim T_3, T_4) = 0.147 + 0.168 = 0.315$$

$$P(T_1, \sim T_3) = 0.315 + 0.042 + 0.032 = 0.389$$

$$P(T_4 | T_1, \sim T_3) = 0.315 / 0.389 = 0.80977$$